

Capacity and Degree-of-Freedom of OFDM Channels with Amplitude Constraint

Saeid Haghighatshoar*, Peter Jung*, Giuseppe Caire*

*Communications and Information Theory Group, Technische Universität Berlin

Emails: saeid.haghighatshoar@tu-berlin.de, peter.jung@tu-berlin.de, caire@tu-berlin.de

Abstract—In this paper, we study the capacity and degree-of-freedom (DoF) scaling for the continuous-time amplitude limited AWGN channels in radio frequency (RF) and intensity modulated optical communication (OC) channels. More precisely, we study how the capacity varies in terms of the OFDM block transmission time T , bandwidth W , amplitude A and the noise spectral density $\frac{N_0}{2}$. We first find suitable discrete encoding spaces for both cases, and prove that they are convex sets that have a semi-definite programming (SDP) representation. Using tools from convex geometry, we find lower and upper bounds on the volume of these encoding sets, which we exploit to drive pretty sharp lower and upper bounds on the capacity. We also study a practical Tone-Reservation (TR) encoding algorithm and prove that its performance can be characterized by the statistical width of an appropriate convex set. Recently, it has been observed that in high-dimensional estimation problems under constraints such as those arisen in Compressed Sensing (CS) statistical width plays a crucial role. We discuss some of the implications of the resulting statistical width on the performance of the TR. We also provide numerical simulations to validate these observations.

Index Terms—Peak-to-Average-Power-Ratio (PAPR), Radio Frequency Channel (RF), Optical Intensity-Modulated Channel (OC), Orthogonal Frequency Division Modulation (OFDM).

1 INTRODUCTION

Shannon in his seminal paper [1] derived the capacity per unit-time of the *continuous-time additive white Gaussian noise* channel (CTAWGN) under input power constraint. The capacity is given by $W \log_2(1 + \text{snr})$, where W is the bandwidth, and where $\text{snr} = \frac{P}{N_0 W}$; P denotes the input power and $\frac{N_0}{2}$ denotes the power spectral density of the white Gaussian noise. He used the sampling theorem for the band-limited signals and the capacity formula $\frac{1}{2} \log_2(1 + \frac{P}{\sigma^2})$ for the discrete-time AWGN (DTAWGN) under the input power constraint P and noise variance σ^2 . A more rigorous proof was later given by Slepian [2], who introduced the Prolate spheroidal wave functions and proved the well-known $2WT$ -result for the dimension of the signal space *essentially* time-limited to T and *essentially* band-limited to W . In brief, the result states that for moderately large *signal-to-noise ratio* (SNR), the capacity scales like $WT \log_2(\text{snr})$, and the *degree-of-freedom* (DoF) for sufficiently high SNR is given by WT .

Although the power-constraint is suitable for theoretical analysis and signal design (due to isometry of the underlying Hilbert spaces), in many practical scenarios in communications systems the front end of the CTAWGN channel is a power amplifier with a limited amplitude A . Thus, it is important to know how the capacity varies under the amplitude constraint. This requires studying the DoF of band-limited signals in $\mathbb{L}_\infty[0, T]$ rather than $\mathbb{L}_2[0, T]$.

For the discrete-time case, the amplitude-limited variant

of the problem was studied by Smith [3], who also showed that the capacity achieving input distribution is discrete and has a finite support. The result was extended to the complex-valued Gaussian channels in [4], and bounds on the capacity were derived in [5]. Recently, tighter bounds were obtained in [6] via a dual capacity result stated in [7]. An interesting lower-bound was given in [8] by using a recent result of Lyubarskii and Vershynin [9] on the existence of tight frames, which in turn uses a deep result of Kashin on the comparison of diameter of certain subsets of Banach spaces under different norms [10]. In brief, the main idea in [8] is to admit some rate loss to transform codewords designed for the power-limited DTAWGN channel (codewords with limited ℓ_2 -norm) into codewords with limited amplitude (limited ℓ_∞ -norm) suitable for amplitude-limited DTAWGN. However, in contrast with the power-limited case, where the discrete-time and the continuous-time variants are related via Hilbert space isometries, the results in [8] does not directly extend to the continuous-time case due to the lack of obvious isometry between $\ell_\infty(\mathbb{Z})$ and $\mathbb{L}_\infty[0, T]$.

In this paper, we study the continuous-time variant of the problem when the transmitter uses an orthonormal collection of waveforms $\mathcal{W}^n = \{\psi_k(t)\}_{k=1}^n$, for signal modulation. The n -dim encoding space for \mathcal{W}^n , called $\mathcal{X}^n \subset \mathbb{C}^n$, consists of all coefficients $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ for which the transmitted waveform $\sum_{k \in [n]} x_k \psi_k(t)$ has a limited amplitude A for all $t \in [0, T]$, where $[n]$ denotes the set of integers $\{1, 2, \dots, n\}$. We focus on the signal space of harmonic waveforms $\mathcal{W}_H^n = \{e^{jk \frac{2\pi}{T} t}\}_{k=1}^n$, known as *orthogonal frequency division modulation* (OFDM). It has been vastly studied in the literature due to its simple implementation (with FFT algorithm), robustness to multi-path fading in wireless scattering channels, and its information theoretic optimality for water-filling type encoding under frequency-selective Gaussian channels [11].

Coding for amplitude-limited OFDM channels is related to the well-known *peak-to-average power ratio* (PAPR) reduction problem. In brief, as the communication is over the AWGN channel, the performance is an increasing function of the average transmit power. However, due to the amplitude limitation A , the peak power is limited. Therefore, the PAPR is a measure of the efficiency of the signal set (code in the signal space), and its minimization corresponds to maximizing the average transmit power. Several techniques have been proposed to tackle this problem such as coding, tone reservation, amplitude clipping, clipping and filtering, tone injection and partial transmit sequences. A summary of important results can be found in the survey papers [12, 13]. There is another collection of work on designing good codes with reasonable PAPR as in the papers [14–17] (see also [18] and the refs. therein). While these codes reduce PAPR, they

also reduce the transmission rate severely, especially for very large number of subcarriers.

Contribution. In this paper, we make a connection between the OFDM for traditional *radio frequency* (RF) channels and OFDM for optical intensity channels (OC) under the intensity constraint A (see [19, 20] and references therein). We show that the encoding set for both cases has a *semi-definite programming* (SDP) representation, which we use to study the capacity scaling performance for both channels. In particular, we show that even though OFDM achieves the optimal DoF scaling WT for OC in high SNR, it seems that, the best possible DoF in moderate SNR is at most λWT , with a multiplicative loss $\lambda \in (0, 1)$. Interestingly, this confirms a recent result of [21], where the authors using the results of Lyubarskii and Vershynin in [9], show (and numerically verify) that if the columns of a subsampled Fourier matrix build a tight frame, then it is possible to make all codewords have a constant PAPR at the cost of a multiplicative rate loss due to those carriers reserved for shaping (*Tone-Reservation* (TR)). We study this problem further, and provide additional evidence that the constant PAPR seems to be achievable, by formulating it as the statistical width of a specific convex set. Our results suggest that, in OFDM systems with a large number of subcarriers, TR is a promising approach in order to achieve an effective PAPR reduction at the cost of a fixed multiplicative loss of DoFs.

2 STATEMENT OF THE PROBLEM

2.1 Basic Setup

Let $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ be a sequence of symbols of length n to be transmitted across a CTAWGN channel. In the OFDM modulation with n subcarriers, the sequence \mathbf{x} is transformed to a base-band signal given by

$$s_b(t) = \sum_{k \in [n]} x_k e^{j2k\pi t}, \quad t \in [0, 1). \quad (1)$$

Here, for convenience, we consider a normalized signal set with a normalized transmission time $T = 1$ and amplitude $A = 1$. We will re-scale the results at the end to make the main system parameters W , T and A and N_0 appear explicitly. The harmonic wave-forms $\mathcal{W}_H = \{e^{j2k\pi t}\}_{k=1}^n$ in (1) form an orthonormal collection with minimum frequency separation 1 over $[0, 1)$, under the inner product $\langle \alpha(t), \beta(t) \rangle = \int_0^1 \alpha(t)^H \beta(t) dt$.

We focus on two main applications of OFDM: 1) in passband modulation for RF channels, and 2) in base-band modulation for OC over intensity modulated channels. In the former, the baseband signal $s_b(t)$ is heterodyned to a sufficiently high *normalized* carrier frequency $f_c \gg n$, and the resulted real-valued passband signal $s(t) = \text{Re}[s_b(t)e^{j2\pi f_c t}]$ is transmitted via the antenna (see Fig. 1). The passband signal $s(t)$ occupies a normalized one-side bandwidth $W = n$ (or more precisely $n+1$) for sufficiently large block-length n . For OC, the real-valued signal $s(t) = 1 + \text{Re}[s_b(t)]$ is used to modulate the light intensity of an optical diode, where we assume that, due to the unipolar nature of the diode, an additional normalized d.c. level of size 1 is added to obtain a positive signal $s(t)$ (see Fig. 1). In both cases, the OFDM symbol \mathbf{x} is detected at the receiver via an electronic circuit and matched filtering. The equivalent discrete-time channel can be modelled as parallel independent complex-

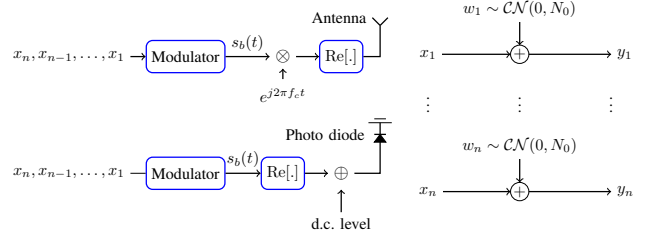


Fig. 1: OFDM for RF and OC channels, and the equivalent discrete-time model.

valued Gaussian channels. This has been shown in Fig. 1, where $\frac{N_0}{2}$ denotes the power spectral density of the channel.

2.2 Behaviour for Large Number of Subcarriers

In this paper, we are interested in the regime where n is very large. Although in the ergodic regime, coding across many consecutive OFDM blocks can be used to further boost the performance, here we mainly focus on the *one-shot* behaviour of the system where an individual OFDM block carries a huge number of symbols. This may be motivated by the recent LTE standards in mobile communications. We are mainly interested to know how the one-shot achievable rate per unit time scales in terms of W , T , A and noise parameter N_0 .

3 ENCODING SPACE

Let $\mathbf{a} : [0, 1) \rightarrow \mathbb{C}^n$ be the vector-valued function given by

$$\mathbf{a}(t) = [e^{-j2\pi t}, e^{-j4\pi t}, \dots, e^{-j2n\pi t}]^T, \quad t \in [0, 1). \quad (2)$$

The baseband signal (1) can be written as $s_b(t) = \mathbf{a}(t)^H \mathbf{x}$. We define the encoding space for RF and OC as

$$\mathcal{X}_{\text{RF}}^n = \{\mathbf{x} \in \mathbb{C}^n : |\mathbf{a}(t)^H \mathbf{x}| \leq 1 \text{ for } t \in [0, 1)\}, \quad (3)$$

$$\mathcal{X}_{\text{OC}}^n = \{\mathbf{x} \in \mathbb{C}^n : 1 + \text{Re}[\mathbf{a}(t)^H \mathbf{x}] \geq 0 \text{ for } t \in [0, 1)\}, \quad (4)$$

where we assume that the amplitude of the RF signal and the d.c. level of the OC signal are both normalized to 1. A code of rate R for communication over AWGN via OFDM signalling is a mapping $\mathcal{C}_{\text{RF}} : [2^{nR}] \rightarrow \mathcal{X}_{\text{RF}}^n$ and $\mathcal{C}_{\text{OC}} : [2^{nR}] \rightarrow \mathcal{X}_{\text{OC}}^n$. Both $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$ are not polyhedral, however, they can be represented as *Linear Matrix Inequalities* (LMI) over the cone of *positive semi-definite* (PSD) matrices, and they have *semi-definite programming* (SDP) representation. We first consider $\mathcal{X}_{\text{RF}}^n$. As a corollary to Theorem 4.24 in [22], we have:

Proposition 3.1: Let $\mathbf{x} \in \mathbb{C}^n$. Then, $\mathbf{x} \in \mathcal{X}_{\text{RF}}^n$ if and only if there exists a Hermitian matrix $\mathbf{H} \in \mathbb{C}^{n \times n}$ such that

$$\begin{bmatrix} \mathbf{H} & \mathbf{x} \\ \mathbf{x}^H & 1 \end{bmatrix} \succeq 0, \quad \sum_{\ell=1}^{n-i} \mathbf{H}_{\ell, \ell+i} = \begin{cases} 1, & i = 0, \\ 0, & i = 1, \dots, n-1. \end{cases} \quad (5)$$

Proof: We only prove one direction by simply applying the Schur's decomposition to the PSD matrix stated in the SDP constraint. Thus, we obtain $\mathbf{H} - \mathbf{x}\mathbf{x}^H \succeq 0$, which implies that for every $\mathbf{w} \in \mathbb{C}^n$, we have $\mathbf{w}^H \mathbf{H} \mathbf{w} \geq |\mathbf{w}^H \mathbf{x}|^2$. Fixing an arbitrary $t \in [0, 1)$ and setting $\mathbf{w} = \mathbf{a}(t) \in \mathbb{C}^n$, we obtain that $\mathbf{w}^H \mathbf{H} \mathbf{w} = 1$, and as a result $|\mathbf{a}(t)^H \mathbf{x}| \leq 1$. Since this is true for every $t \in [0, 1)$, the proof is complete. ■

We can obtain a similar representation for the set $\mathcal{X}_{\text{OC}}^n$. We need some notations first. We define \mathcal{S}_+ as the set of all $(n+1) \times (n+1)$ Hermitian PSD matrices, and $\mathcal{S}_+^\circ = \{\mathbf{X} \in \mathcal{S}_+ : \text{tr}[\mathbf{X}] = 1\}$ as the affine subspace of \mathcal{S}_+ with unit trace. Note that \mathcal{S}_+° is a closed convex subset of \mathcal{S}_+ . Moreover,

for any $\mathbf{X} \in \mathcal{S}_+^\circ$, we have $\text{tr}[\mathbf{X}] = \sum_{i \in [n+1]} \lambda_i = 1$, where $\lambda_i \geq 0$ denote the eigen-values of \mathbf{X} . This implies that \mathcal{S}_+° is also bounded. We define $\text{sud} : \mathcal{S}_+^\circ \rightarrow \mathbb{C}^n$ as the sum-diagonal map, which for every $\mathbf{X} \in \mathcal{S}_+^\circ$ and for $k \in [n]$, gives a vector $\text{sud}(\mathbf{X})$ by

$$[\text{sud}(\mathbf{X})]_k = 2 \text{diag}(\mathbf{X}, -k) := \sum_{\ell=k}^{n+1} 2 \mathbf{X}_{\ell, \ell-k+1}. \quad (6)$$

As a corollary to Theorem 3 in [23], we have the following representation of $\mathcal{X}_{\text{OC}}^n$.

Proposition 3.2: Let $\mathbf{x} \in \mathbb{C}^n$. Then, $\mathbf{x} \in \mathcal{X}_{\text{OC}}^n$ if and only if there is an $\mathbf{X} \in \mathcal{S}_+^\circ$ such that $\mathbf{x} = \text{sud}(\mathbf{X})$. \square

Proof: One side is again easy to prove. Let $t \in [0, 1]$ and let $\mathbf{b}(t) = [1, e^{j2\pi t}, e^{j4\pi t}, \dots, e^{j2n\pi t}]^T$. It is not difficult to see that $\mathbf{b}(t)\mathbf{b}(t)^H$ is a Toeplitz matrix with 1 on its main diagonal and with $e^{j2k\pi t}$ and $e^{-j2k\pi t}$ on its k -th lower and k -th upper diagonal respectively. Now let $\mathbf{x} \in \mathbb{C}^n$ and suppose there is an $\mathbf{X} \in \mathcal{S}_+^\circ$, with $\mathbf{x} = \text{sud}(\mathbf{X})$. This implies that

$$0 \leq \mathbf{b}(t)^H \mathbf{X} \mathbf{b}(t) = \text{tr}[\mathbf{X} \mathbf{b}(t) \mathbf{b}(t)^H] \quad (7)$$

$$= 1 + \sum_{k \in [n]} \text{diag}(\mathbf{X}, -k) e^{j2k\pi t} + \text{diag}(\mathbf{X}, k) e^{-j2k\pi t} \quad (8)$$

$$= 1 + \text{Re}[\mathbf{a}(t)^H \text{sud}(\mathbf{X})] = 1 + \text{Re}[\mathbf{a}(t)^H \mathbf{x}]. \quad (9)$$

Since this is true for every $t \in [0, 1]$, then $\mathbf{x} \in \mathcal{X}_{\text{OC}}^n$. \blacksquare

The next proposition shows some of the properties and also the relation between $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$.

Proposition 3.3: The sets $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$ are compact and convex subsets of \mathbb{C}^n . Moreover, $\mathcal{X}_{\text{RF}}^n = \bigcap_{\phi \in [0, 2\pi)} \{e^{j\phi} \mathcal{X}_{\text{OC}}^n\}$. In particular, $\mathcal{X}_{\text{RF}}^n \subset \mathcal{X}_{\text{OC}}^n$. \square

Proof: The compactness and the convexity can be directly checked from the definition. However, it is also seen from Proposition 3.1 and 3.2, that $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$ are obtained from linear projection of compact and convex subsets of $(n+1) \times (n+1)$ PSD matrices, thus, they must be compact and convex.

To prove the next part, let $\mathcal{L} = \bigcap_{\phi \in [0, 2\pi)} \{e^{j\phi} \mathcal{X}_{\text{OC}}^n\}$. Note that \mathcal{L} is a symmetric set, i.e., $e^{j\phi} \mathcal{L} = \mathcal{L}$ for all $\phi \in [0, 2\pi)$. It is not also difficult to see that it is the largest symmetric set contained in $\mathcal{X}_{\text{OC}}^n$. We use this property to prove the theorem.

First note that since $\mathcal{X}_{\text{RF}}^n$ is itself a symmetric set, it must be contained in \mathcal{L} , i.e., $\mathcal{X}_{\text{RF}}^n \subset \mathcal{L}$. To prove the other direction, let $\mathbf{x} \in \mathcal{L}$ and let $t \in [0, 1]$. From the symmetry of \mathcal{L} , it results that for every $\phi \in [0, 2\pi)$, and in particular, $\phi = \pi - \angle \mathbf{a}(t)^H \mathbf{x}$, the vector $e^{j\phi} \mathbf{x}$ belongs to \mathcal{L} , thus, to $\mathcal{X}_{\text{OC}}^n$. Hence, we must have $1 + \text{Re}[e^{j\phi} \mathbf{a}(t)^H \mathbf{x}] \geq 0$, which implies that $|\mathbf{a}(t)^H \mathbf{x}| \leq 1$. Since this is true for an arbitrary $t \in [0, 1]$, we obtain $\mathcal{L} \subset \mathcal{X}_{\text{RF}}^n$. This completes the proof. \blacksquare

4 LOWER AND UPPER BOUND ON THE VOLUME

In this section, we derive lower and upper bounds on the volume of $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$ via results in convex geometry.

4.1 Real Embedding

Since the results can be stated more conveniently in the real-valued case we first embed both sets $\mathcal{X}_{\text{RF}}^n$ and $\mathcal{X}_{\text{OC}}^n$ in $\mathbb{R}^{n \times 2}$. We identify every $\mathbf{x} \in \mathbb{C}^n$ by the $n \times 2$ matrix $X = \mathcal{R}(\mathbf{x}) = [\text{Re}[\mathbf{x}], \text{Im}[\mathbf{x}]]$, where $\mathcal{R} : \mathbb{C}^n \rightarrow \mathbb{R}^{n \times 2}$ denotes this real embedding. We denote the inverse map by $\mathcal{I} : \mathbb{R}^{n \times 2} \rightarrow \mathbb{C}^n$. It is not difficult to check the isometry

$$\langle \mathcal{R}(\mathbf{x}_1), \mathcal{R}(\mathbf{x}_2) \rangle = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle_{\mathbb{R}} = \text{Re}[\mathbf{x}_1^H \mathbf{x}_2], \quad (10)$$

where the first inner product is the conventional inner product between matrices in $\mathbb{R}^{n \times 2}$ given by $\langle X_1, X_2 \rangle = \text{tr}[X_1^T X_2]$. We define $A(t) = \mathcal{R}(\mathbf{a}(-t))$, where $\mathbf{a}(t)$ is given by (2). We can check that the columns of $A(t)$ are given by the vector $\mathbf{c}(t) := [\cos(2\pi t), \cos(4\pi t), \dots, \cos(2n\pi t)]^T$ and $\mathbf{s}(t) := [\sin(2\pi t), \sin(4\pi t), \dots, \sin(2n\pi t)]^T$. We also define the matrix $B(t) := [-\mathbf{s}(t), \mathbf{c}(t)]$. We can also see that multiplying the vector $\mathbf{a}(t)$ by the constant phase $e^{j\phi}$ is equivalent to rotating the columns of $A(t)$ by a one parameter rotation group. More precisely, we have

$$\mathcal{R}(e^{j\phi} \mathbf{a}(-t)) = \cos(\phi) A(t) + \sin(\phi) B(t). \quad (11)$$

4.2 Polar of a set

For a subset \mathcal{C} of $\mathbb{R}^{n \times 2}$, we define the polar (symmetric polar) of \mathcal{C} as

$$\mathcal{C}^\circ = \{Y \in \mathbb{R}^{n \times 2} : \sup_{X \in \mathcal{C}} |\langle X, Y \rangle| \leq 1\}. \quad (12)$$

The polar is typically defined without the absolute value but since we always work with symmetric sets, we keep the absolute value. It is not difficult to see from (12) that \mathcal{C}° is always closed and convex, and contains the origin. From duality in convex geometry, it results that if \mathcal{C} is convex, closed and symmetric then $\mathcal{C}^{\circ\circ} = \mathcal{C}$. From (12) it is not also difficult to see that the polar set does not change if we replace \mathcal{C} by $\text{cl}\{\text{co}\{\mathcal{C}, -\mathcal{C}\}\}$, where cl and co denote the closure and convex hull operation. This implies that for every set \mathcal{C} , we have $\mathcal{C}^{\circ\circ} = \text{cl}\{\text{co}\{\mathcal{C}, -\mathcal{C}\}\}$. Another property that we will use is that if $\mathcal{C} \subset \mathcal{B}$ then $\mathcal{B}^\circ \subset \mathcal{C}^\circ$, and $(\lambda \mathcal{C})^\circ = \frac{1}{\lambda} \mathcal{C}^\circ$.

4.3 Mahler and Bottleneck Conjecture

Let $\mathcal{K} \subset \mathbb{R}^{n \times 2}$ be a symmetric convex set with a polar set \mathcal{K}° . The Mahler volume of \mathcal{K} is defined as $\nu(\mathcal{K}) = \text{vol}(\mathcal{K}) \text{vol}(\mathcal{K}^\circ)$, where vol denotes the volume. Mahler in [24] conjectured that for every such \mathcal{K}

$$\frac{4^{2n}}{(2n)!} \leq \nu(\mathcal{K}) \leq \frac{\pi^{2n}}{(n!)^2}, \quad (13)$$

where the upper and the lower bound are achieved for the sphere and the cube respectively (note that in our case the dimension is $2n$, and the conjecture is stated for $2n$). The upper bound was proven by Santaló [25], and is known as the Blaschke-Santaló inequality. The conjecture for the lower bound is still open but a weak variant of it, known as the *Bottleneck conjecture*, has been proven which implies the Mahler conjecture up to a factor of $(\frac{\pi}{4})^{2n} \gamma_n$, where γ_n is monotonic factor that begins at $\frac{4}{\pi}$ and increases to $\sqrt{2}$ as the n goes to infinity [26]. In our case, it gives the lower bound $\nu(\mathcal{K}) \geq \sqrt{2} \frac{\pi^{2n}}{(2n)!}$ for any symmetric convex set \mathcal{K} .

4.4 Lower bound on $\text{vol}(\mathcal{X}_{\text{OC}}^n)$

Let $\mathcal{X}_{\text{OC}}^n$ be encoding set defined by (4). We define the symmetric part of $\mathcal{X}_{\text{OC}}^n$ as $\overline{\mathcal{X}_{\text{OC}}^n} = (-\mathcal{X}_{\text{OC}}^n \cap \mathcal{X}_{\text{OC}}^n)$, where it is easy to see that

$$\overline{\mathcal{X}_{\text{OC}}^n} = \{\mathbf{x} \in \mathbb{C}^n : -1 \leq \text{Re}[\mathbf{a}(t)^H \mathbf{x}] \leq 1 \text{ for } t \in [0, 1]\}.$$

Let $\mathcal{K}_1 \subset \mathbb{R}^{n \times 2}$ be the real embedding of $\overline{\mathcal{X}_{\text{OC}}^n}$. It is not difficult to see that \mathcal{K}_1 can be identified with the polar of the set $\mathcal{A}_\pm = \bigcup_{t \in [0, 1]} \{A(t), -A(t)\}$. In particular, we have that $\mathcal{K}_1^\circ = \text{cl}(\text{co}(\mathcal{A}_\pm))$. We also define $\mathcal{A}_+ = \bigcup_{t \in [0, 1]} \{A(t)\}$, $\mathcal{A} = \text{co}(\mathcal{A}_+)$, and $\mathcal{E} = \mathcal{I}(\mathcal{A})$. It is not difficult to check that the convexity remains invariant under the real embedding \mathcal{R} and its inverse \mathcal{I} . This implies that

$\mathcal{E} = \text{co}(\cup_{t \in [0,1)} \{\mathbf{a}(-t)\}) = \text{co}(\cup_{t \in [0,1)} \{\mathbf{a}(t)\})$ due to the periodicity of $\mathbf{a}(t)$ in $[0, 1)$.

We first prove that $\mathcal{K}_1^\circ \subset (\mathcal{A} - \mathcal{A})$, where $(\mathcal{A} - \mathcal{A}) := \{X - Y : X, Y \in \mathcal{A}\}$ denotes the Minkowski difference of \mathcal{A} . We need the following preliminary lemma.

Lemma 4.1: The set \mathcal{A} contains the origin.

Proof: Note that $A(t) \in \mathcal{A}$ for every $t \in [0, 1)$. Let μ be the uniform probability measure over $[0, 1)$. Since \mathcal{A} is convex, it results that $\mathbf{0} = \int_0^1 A(t) \mu(dt) \in \mathcal{A}$, thus, \mathcal{A} contains the origin. ■

Proposition 4.2: Let \mathcal{A} and \mathcal{K}_1° be as defined before. Then $\mathcal{K}_1^\circ \subset (\mathcal{A} - \mathcal{A})$. □

Proof: From Lemma 4.1, it results that the convex set \mathcal{A} contains the origin. This implies that $\mathcal{A}, -\mathcal{A} \subset (\mathcal{A} - \mathcal{A})$, and especially $\mathcal{A}_\pm \subset (\mathcal{A} - \mathcal{A})$. Since $(\mathcal{A} - \mathcal{A})$ is convex, it results that $\mathcal{K}_1^\circ = \text{cl}(\text{co}(\mathcal{A}_\pm)) \subset (\mathcal{A} - \mathcal{A})$. ■

Consider the set \mathcal{E} defined before, where it is seen that \mathcal{E} is the convex hull of the periodic curve $\cup_{t \in [0,1)} \{\mathbf{a}(t)\}$. Schönberg in [27] using isoperimetric inequalities proved that $\text{vol}(\mathcal{E})$ is strictly larger than the volume of convex hull of any other periodic curve with the same length, where the equality holds if and only if the curve is obtained by some linear isometry from the curve $\cup_{t \in [0,1)} \{\mathbf{a}(t)\}$. He also showed that $\text{vol}(\mathcal{E}) = \frac{2^n \pi^n n!}{(2n)!}$. Thus, we obtain that $\text{vol}(\mathcal{A}) = \text{vol}(\mathcal{E}) = \frac{2^n \pi^n n!}{(2n)!}$. Moreover, from Rogers-Shephard inequality [28], $\text{vol}(\mathcal{A} - \mathcal{A})$ can be upper bounded by $\binom{4n}{2n} \text{vol}(\mathcal{A})$. Thus, using Proposition 4.2, we obtain

$$\text{vol}(\mathcal{K}_1^\circ) \leq \text{vol}(\mathcal{A} - \mathcal{A}) \leq \binom{4n}{2n} \frac{2^n \pi^n n!}{(2n)!}. \quad (14)$$

Since \mathcal{K}_1 and \mathcal{K}_1° are symmetric and convex, applying the result of Bottleneck conjecture, i.e., $\nu(\mathcal{K}_1) = \text{vol}(\mathcal{K}_1) \text{vol}(\mathcal{K}_1^\circ) \geq \sqrt{2} \frac{\pi^{2n}}{(2n)!}$, and using (14), we obtain that

$$\text{vol}(\mathcal{X}_{\text{OC}}^n) \geq \text{vol}(\mathcal{K}_1) \geq \frac{\sqrt{2} \pi^n}{2^n n! \binom{4n}{2n}}. \quad (15)$$

4.5 Lower Bound on $\text{vol}(\mathcal{X}_{\text{RF}}^n)$

Let \mathcal{K}_2 be the real embedding of $\mathcal{X}_{\text{RF}}^n$. Note that from the definition of $\mathcal{X}_{\text{RF}}^n$ in (4) and the one-dimensional rotation property mentioned in (11), we can see that \mathcal{K}_2 is given by the polar of the parametric set

$$\mathcal{A}_\phi := \cup_{t \in [0,1), \phi \in [0, 2\pi)} \{\cos(\phi)A(t) + \sin(\phi)B(t)\}, \quad (16)$$

where in the special case $\phi \in \{0, \pi\}$, we obtain $\mathcal{A}_\pm = \cup_{t \in [0,1)} \{A(t), -A(t)\}$. Let $\mathcal{B}_\pm = \cup_{t \in [0,1)} \{B(t), -B(t)\}$. It is not difficult to see that finding a lower bound for $\text{vol}(\mathcal{K}_2)$ at least requires finding an upper bound on the volume of the convex hull of $\mathcal{A}_\pm \cup \mathcal{B}_\pm$. This seems quite challenging especially that, in contrary to Rogers-Shephard inequality, there is no universal upper bound on the volume of the convex hull of the union of two different sets in terms of volume of their individual convex hulls. Instead, we use another approach to find a lower bound on $\text{vol}(\mathcal{X}_{\text{RF}}^n)$. We consider two grid of size n on $[0, 1)$ given by $\mathcal{G}_1 = \{\frac{i-1}{n} : i \in [n]\}$ and $\mathcal{G}_2 = \{\frac{2i-1}{2n} : i \in [n]\}$. We set \mathbf{G}_1 as a matrix whose columns are given by $\mathbf{a}(g)$, $g \in \mathcal{G}_1$. We define \mathbf{G}_2 similarly by using grid points from \mathcal{G}_2 . It is not difficult to see that \mathbf{G}_1 coincides with the DFT matrix. In particular, $\mathbf{G}_1^H \mathbf{G}_1 = \mathbf{G}_2^H \mathbf{G}_2 = n \mathbf{I}_n$. We also define $\mathbf{K} = \frac{1}{n} \mathbf{G}_2^H \mathbf{G}_1$. We can simply check that \mathbf{K} is a Toeplitz matrix with $\mathbf{K}_{\ell m} = K(\ell - m + 1/2)$ where $K(t) = e^{j\pi t \frac{\sin(n\pi t)}{n \sin(\pi t)}}$. Note that $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ is a uniform

grid over $[0, 1)$ of size $2n$, with an oversampling factor 2. Let $\bar{\mathcal{X}}_{\text{RF}}^n$ be the approximation of $\mathcal{X}_{\text{RF}}^n$ via grid \mathcal{G} , i.e., $\bar{\mathcal{X}}_{\text{RF}}^n = \{\mathbf{x} : \max_{g \in \mathcal{G}} |\mathbf{a}(g)^H \mathbf{x}| \leq 1\}$. In [29], it was shown that for an oversampling factor $\gamma > 1$

$$\sup_{t \in [0,1)} |\mathbf{a}(t)^H \mathbf{x}| \leq \frac{1}{\cos(\frac{\pi}{2\gamma})} \max_{g \in \mathcal{G}} |\mathbf{a}(g)^H \mathbf{x}|. \quad (17)$$

In our case $\gamma = 2$, and we obtain that $\mathcal{X}_{\text{RF}}^n \subset \bar{\mathcal{X}}_{\text{RF}}^n \subset \sqrt{2} \mathcal{X}_{\text{RF}}^n$, which implies that $\text{vol}(\mathcal{X}_{\text{RF}}^n) \geq 2^{-n} \text{vol}(\bar{\mathcal{X}}_{\text{RF}}^n)$.

Let $\mathcal{U} = \{\mathbf{x} \in \mathbb{C}^n : \max_{k \in [n]} |x_k| \leq 1\}$ be the complex ℓ_∞ -ball. Note that $\bar{\mathcal{X}}_{\text{RF}}^n = \mathcal{C}_1 \cap \mathcal{C}_2$, where $\mathcal{C}_i = \{\mathbf{x} : \mathbf{G}_i^H \mathbf{x} \in \mathcal{U}\}$, $i \in [2]$. To lower bound $\text{vol}(\bar{\mathcal{X}}_{\text{RF}}^n)$, we need to find a lower bound for the intersection of two convex sets \mathcal{C}_1 and \mathcal{C}_2 . Since both \mathcal{C}_1 and \mathcal{C}_2 are symmetric sets containing the origin, we find a constant $\beta_n \geq 1$ such that $\frac{1}{\beta_n} \mathcal{C}_1 \subset \mathcal{C}_2$, which implies that $\frac{1}{\beta_n} \mathcal{C}_1 \subset \mathcal{C}_1 \cap \mathcal{C}_2$. This has been pictorially shown in Fig. 2.

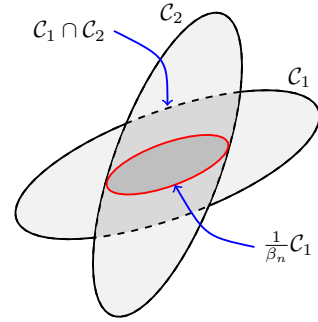


Fig. 2: Approximating $\mathcal{C}_1 \cap \mathcal{C}_2$ by a scaled version of \mathcal{C}_1 .

Proposition 4.3: There is a parameter $\beta_n \geq 1$ such that $\frac{1}{\beta_n} \mathcal{C}_1 \subset \mathcal{C}_1 \cap \mathcal{C}_2$. Moreover, $\beta_n \approx \theta \log_2(n)$ for $\theta \approx \frac{\log(2)}{\pi}$ for sufficiently large n . □

Proof: From the definition of \mathcal{C}_1 and \mathcal{C}_2 , it is seen that the required scaling factor β_n is given by

$$\begin{aligned} \beta_n &= \max_{\mathbf{x} \in \mathcal{U}} \|\mathbf{G}_2^H \mathbf{G}_1 \mathbf{x}\|_\infty = \max_{\mathbf{x} \in \mathcal{U}} \|\mathbf{K} \mathbf{x}\|_\infty \\ &\leq \sum_{\ell \in [n]} |\mathbf{K}(\ell + 1/2)| \leq \sum_{\ell \in [n]} \frac{2}{\pi(2\ell + 1)} \approx \theta \log_2(n), \end{aligned} \quad (18)$$

where $\theta \approx \frac{\log(2)}{\pi}$. This completes the proof. ■

Proposition 4.4: For sufficiently large n , the volume of $\mathcal{X}_{\text{RF}}^n$ is lower bounded by $\frac{\pi^n}{2^n \theta^{2n} \log_2(n)^{2n} n^n}$. □

Proof: First note that the volume of ℓ_∞ -ball \mathcal{U} is given by π^n . Since $\frac{1}{\sqrt{n}} \mathbf{G}_i$ are unitary matrices, we have $\text{vol}(\mathcal{C}_i) = \frac{\text{vol}(\mathcal{U})}{(\sqrt{n})^{2n}} = \frac{\pi^n}{n^n}$. Using Proposition 4.3, we obtain

$$\text{vol}(\mathcal{X}_{\text{RF}}^n) \geq 2^{-n} \text{vol}(\bar{\mathcal{X}}_{\text{RF}}^n) \geq \frac{\text{vol}(\mathcal{C}_1)}{2^n \beta_n^{2n}} = \frac{\pi^n}{2^n \theta^{2n} \log_2(n)^{2n} n^n},$$

which is the desired result. ■

4.6 Upper Bound on the Volume

In [19], it was shown that the encoding set $\mathcal{X}_{\text{OC}}^n$ is contained in $\mathcal{E} = \text{co}(\cup_{t \in [0,1)} \{\mathbf{a}(t)\})$ as defined before. This can be written as

$$\mathcal{E} = \{\mathbf{x} \in \mathbb{C}^n : \exists \mu \text{ such that } \int_0^1 \mathbf{a}(t) \mu(dt) = \mathbf{x}\}, \quad (19)$$

where μ denotes a probability measure. This can be easily proved since if $\mathbf{x} \in \mathcal{X}_{\text{OC}}^n$, then $x(t) = 1 + \text{Re}[\mathbf{a}(t)^H \mathbf{x}]$ is positive with $\int_0^1 x(t) dt = 1$. And it can be easily checked that $\mu(dt) = x(t) dt$ as a probability measure over $[0, 1)$

satisfies $\int_0^1 \mathbf{a}(t)\mu(dt) = \mathbf{x}$, thus, \mathbf{x} should belong to \mathcal{E} . As we explained in Section 4, the volume of \mathcal{E} was computed by Schönberg in [27] to be $\text{vol}(\mathcal{E}) = \frac{2^n \pi^n n!}{(2n)!}$. Hence, from $\mathcal{X}_{\text{RF}}^n \subset \mathcal{X}_{\text{OC}}^n \subset \mathcal{E}$, we have that

$$\text{vol}(\mathcal{X}_{\text{RF}}^n) \leq \text{vol}(\mathcal{X}_{\text{OC}}^n) \leq \frac{2^n \pi^n n!}{(2n)!}. \quad (20)$$

5 CAPACITY RESULTS

5.1 Lower Bounds on the Capacity

We define $\mathcal{P}_{\mathcal{X}_{\text{RF}}^n} = \{p : \text{supp}(p) \subset \mathcal{X}_{\text{RF}}^n\}$, as the set of all probability measures supported on $\mathcal{X}_{\text{RF}}^n$. We define $\mathcal{P}_{\mathcal{X}_{\text{OC}}^n}$ similarly. Since $\mathcal{X}_{\text{RF}}^n \subset \mathcal{X}_{\text{OC}}^n$, we have $\mathcal{P}_{\mathcal{X}_{\text{RF}}^n} \subset \mathcal{P}_{\mathcal{X}_{\text{OC}}^n}$. Let $p \in \mathcal{P}_{\mathcal{X}_{\text{RF}}^n}$, and let $\mathbf{x} \sim p$. We define the mutual information between the input and output of the channel by

$$I_p(\mathbf{x}; \mathbf{y}) = h_p(\mathbf{y}) - h_p(\mathbf{y}|\mathbf{x}) = h_p(\mathbf{y}) - h(\mathbf{w}), \quad (21)$$

where $h_p(\cdot)$ denotes the differential entropy under $\mathbf{x} \sim p$, and where \mathbf{w} is the additive noise vector (see Fig. 1). Note that since $\mathcal{X}_{\text{RF}}^n$ is a compact set, $\mathbf{y} = \mathbf{x} + \mathbf{w}$ has a well-defined covariance matrix $\Sigma_y = \Sigma_x + N_0 \mathbf{I}_n$, thus,

$$h_p(\mathbf{y}) \leq h(\mathbf{y}_g) = \log_2((\pi e)^n |\Sigma_y|) < \infty, \quad (22)$$

where \mathbf{y}_g is a complex Gaussian vector with the same covariance as Σ_y . This implies that $I_p(\mathbf{x}; \mathbf{y})$ is always well-defined for every $p \in \mathcal{P}_{\mathcal{X}_{\text{RF}}^n}$. We also define

$$C_{\mathcal{X}_{\text{RF}}^n}^{(n)} = \sup_{p \in \mathcal{P}_{\mathcal{X}_{\text{RF}}^n}} I_p(\mathbf{x}; \mathbf{y}). \quad (23)$$

Proposition 5.1: Let $C_{\mathcal{X}_{\text{RF}}^n}^{(n)}$ be defined as in (23). Then, we have $C_{\mathcal{X}_{\text{RF}}^n}^{(n)} \geq \log_2(\text{vol}(\mathcal{X}_{\text{RF}}^n)) - \log_2((\pi e N_0)^n)$. \square

Proof: Let p_u be the uniform probability distribution on $\mathcal{X}_{\text{RF}}^n$. Since $\mathcal{X}_{\text{RF}}^n$ is compact p_u is well-defined. Moreover, $C_{\mathcal{X}_{\text{RF}}^n}^{(n)} \geq I_{p_u}(\mathbf{x}; \mathbf{y}) = h_{p_u}(\mathbf{y}) - h(\mathbf{w}) \geq h_{p_u}(\mathbf{y}|\mathbf{w}) - h(\mathbf{w}) = h_{p_u}(\mathbf{x}) - h(\mathbf{w}) = \log_2(\text{vol}(\mathcal{X}_{\text{RF}}^n)) - \log_2((\pi e N_0)^n)$.

This completes the proof. \blacksquare

A result similar to Proposition 5.1 holds for $\mathcal{X}_{\text{OC}}^n$. To state the lower bound in terms of the physical parameters including transmission time T , bandwidth W , amplitude A , and noise parameter N_0 , we need to normalize transmission time by T , bandwidth by W , take $n = WT$, scale the harmonic basis functions $\psi_k(t) = e^{j2k\pi t}$, $t \in [0, 1]$ by $\frac{1}{\sqrt{T}}\psi_k(\frac{t}{T})$, $t \in [0, T]$, normalize the amplitude of the signal by $A\sqrt{T}$, and keep the noise parameter the same as N_0 . Applying this normalization, and using $n! \approx \sqrt{2\pi n}(n/e)^n$, for sufficiently large $n = WT$, we obtain the following lower bounds for the *capacity per unit time* of RF and OC channels:

$$C_{\text{OC}}(W) \geq W \left(\log_2\left(\frac{A^2}{2WN_0}\right) - 4 \right) \quad (24)$$

$$C_{\text{RF}}(W) \geq W \left(\log_2\left(\frac{A^2}{2WN_0}\right) + 3 - 2\log_2(\alpha) \right). \quad (25)$$

The loss 4 in (24) mainly results from Rogers-Shephard inequality that $\text{vol}(\mathcal{A} - \mathcal{A}) \leq \binom{4n}{2n} \text{vol}(\mathcal{A})$ and can be further improved. At least, it can be reduced to $3.03 = 4 - 2\log_2(4/\pi)$ if the Mahler conjecture is true.

We have the additive term $3 \approx \log_2(\frac{1}{e\theta^2})$ in (25) but the best value that we could find for α in (25) is given by $\log_2(WT)$, which results from the lower bound in Proposition 4.4. It seems that finding a universal α independent of WT may not be possible. This suggests that the true be-

haviour of the capacity of the amplitude-limited RF channel in the *one-shot* regime is of the form $\lambda W \log_2(\frac{A^2}{2WN_0})$ for some $\lambda \in (0, 1)$, thus, indicating that a loss of DoFs with respect to its power-limited counterpart is unavoidable. This would also suggest that an effective and more practical way to approach the *one-shot* capacity of this channel consists in fixing WT to some sufficiently large value, and reserving a fixed fraction of subcarriers to keep the signal's PAPR under control. This approach, known as Tone Reservation (TR), has been widely investigated in the literature (see refs. in Section 1) and will be treated in a novel way in Section 6, by exploiting our SDP characterization of the set $\mathcal{X}_{\text{RF}}^n$.

5.2 Upper Bound on the Capacity

Using the results in [19] and upper bounds on the volume in (20) derived in Section 4, we obtain an upper bound on the capacity per unit time of OC in high-SNR regime, which also gives an upper bound for the capacity of RF. After suitable scaling we have

$$C_{\text{RF}}(W) \leq C_{\text{OC}}(W) \leq W \log_2\left(\frac{A^2}{2WN_0}\right), \quad (26)$$

which shows that the lower bounds in (24) is tight for OC up to a finite loss in SNR.

6 TONE RESERVATION AND THE PAPR PROBLEM

In this section, we investigate tone-reservation algorithm (TR) for an individual OFDM block of large dimension $n = WT$. Of course, in practice, coding across a sufficiently large number of OFDM blocks can be used to further increase the reliability. From capacity result in (25), it seems that the loss in SNR given by α might scale as $\log_2(WT)$. Thus, to compensate the loss in SNR, it might be necessary to encode over blocks with smaller WT . For a fixed bandwidth W , it implies that the OFDM packets should be made smaller, and coding over consecutive blocks should be used to achieve the optimal scaling $W \log_2(\frac{A^2}{2WN_0})$.

Recently, Ilic and Strohmer [21] studied the performance of Tone Reservation (TR) using the results of Lyubarskii and Vershynin in [9]. Although not yet rigorous, their results suggest that a constant PAPR for all codewords is possible provided that a fixed fraction of carriers are devoted to waveform shaping, and PAPR reduction. This shows that, the DoF WT up to a multiplicative loss seems to be achievable. In this section, using the SDP representation for $\mathcal{X}_{\text{RF}}^n$, we prove that the best PAPR for TR is given by the statistical width of an appropriate convex set that we define.

In TR an OFDM block of length n is divided into two sub-blocks: a block $\mathcal{I} \subset [n]$ of size $m = \lambda n$, $\lambda \in (0, 1)$, containing the information symbols, and a block $\mathcal{R} = [n] \setminus \mathcal{I}$ containing the symbols used for waveform shaping to reduce the PAPR. This reduces the rate by a factor λ . Moreover, an extra power is transmitted for the redundant symbols. Let $\mathbf{s} \in \mathbb{C}^m$ be the sequence of information symbols. We assume that each component of \mathbf{s} is selected from a given signal constellation such as QAM. Using the SDP representation of $\mathcal{X}_{\text{RF}}^n$ in (5), and after suitable normalization, we obtain the following SDP for the optimal selection of redundant symbols:

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathbb{C}^n} \text{tr}[\mathbf{H}] \text{ subject to} \quad (27)$$

$$\begin{bmatrix} \mathbf{H} & \mathbf{x} \\ \mathbf{x}^H & 1 \end{bmatrix} \succeq 0, \sum_{\ell=1}^{n-i} \mathbf{H}_{\ell, \ell+i} = 0, i \in [n-1], \mathbf{x}_{\mathcal{I}} = \mathbf{s},$$

where $\mathbf{x}_{\mathcal{I}}$ is a sub-vector of \mathbf{x} containing the components in location \mathcal{I} . It is not difficult to check that the minimum power loss for transmitting \mathbf{s} is given by $\ell(\mathbf{s}, \mathcal{I}) := \frac{\text{tr}[\mathbf{H}^*]}{\|\mathbf{s}\|^2}$, where \mathbf{H}^* is the optimal matrix obtained from (27). We suppose that each symbols s_i in \mathbf{s} is generated i.i.d. from a given distribution p_s , where $\mathbb{E}[s_i] = 0$ and $\mathbb{E}[|s_i|^2] = 1$. For large n , we have that $\ell(\mathbf{s}, \mathcal{I}) \approx \text{tr}[\mathbf{H}^*]/m$. Note that $\ell(\mathbf{s}, \mathcal{I})$ is a random variable depending on the information symbols \mathbf{s} and their location \mathcal{I} .

Let $\mathcal{C} \subset \mathbb{C}^m$ be a convex set. We define the statistical width of \mathcal{C} under i.i.d. sampling induced by distribution $p_s(\mathbf{s})$ by $\omega(\mathbf{s}, \mathcal{C}) := \sup_{\mathbf{u} \in \mathcal{C}} |\langle \mathbf{u}, \mathbf{s} \rangle_{\mathbb{R}}|^2$, where $\langle \mathbf{u}, \mathbf{s} \rangle_{\mathbb{R}} = \text{Re}[\mathbf{u}^H \mathbf{s}]$, and its average by $\bar{\omega}(\mathcal{C}) = \mathbb{E}_s[\omega(\mathbf{s}, \mathcal{C})]$. When p_s is a Gaussian distribution, this is known as the squared Gaussian width of the set \mathcal{C} , which plays a crucial role in characterizing the minimum number of measurements in high-dimensional estimation under constraints such as Compressed Sensing [30, 31]. For a given position of information symbols \mathcal{I} , let us define

$$\mathcal{C}_{\mathcal{I}} = \{\mathbf{u} \in \mathbb{C}^m : \exists \mathbf{y} \text{ s.t. } \begin{bmatrix} \mathbb{T}[1; \mathbf{y}] & \mathbf{z}(\mathbf{u}) \\ \mathbf{z}(\mathbf{u})^H & 1 \end{bmatrix} \succeq 0\}, \quad (28)$$

where $\mathbb{T}[1; \mathbf{y}]$ denotes a Hermitian Toeplitz matrix with first column $[1; \mathbf{y}]$, and where $\mathbf{z}(\mathbf{u}) \in \mathbb{C}^n$ is a vector containing \mathbf{u} in indices corresponding to \mathcal{I} and zero elsewhere. It is not difficult to check that $\mathcal{C}_{\mathcal{I}}$ is a convex set containing the origin. Then, we obtain the following proposition.

Proposition 6.1: Let $\ell(\mathbf{s}, \mathcal{I})$ be the power-loss (PAPR) as defined before. Then, $\ell(\mathbf{s}, \mathcal{I}) = \omega(\mathbf{s}, \mathcal{C}_{\mathcal{I}})/m$, where $\omega(\mathbf{s}, \mathcal{C}_{\mathcal{I}})$ is the width of $\mathcal{C}_{\mathcal{I}}$ for the \mathbf{s} with i.i.d. components $\sim p_s$. \square

Proof: We use the duality to prove the result. First note that the Slater's condition [32] holds for the SDP constraint because for sufficiently large $\tau \in \mathbb{R}_+$, by setting $\mathbf{H} = \tau \mathbf{I}_n$, the matrix $\begin{bmatrix} \mathbf{H} & \mathbf{x} \\ \mathbf{x}^H & 1 \end{bmatrix}$ will satisfy the constraints and will be strictly positive, i.e., will lie in the relative interior of the convex constrained set. Thus, we have the strong duality without any duality gap. Introducing dual variables $\mathbf{z} \in \mathbb{C}^n$ for the constraint $\mathbf{x}_{\mathcal{I}} = \mathbf{s}$, $\mathbf{y} \in \mathbb{C}^{n-1}$ for the constraints in \mathbf{H} , and w for the constraint of having 1 at the intersection of the last row and last column, we obtain the dual of the SDP in (27) as follows

$$\sup_{\mathbf{z} \in \mathbb{C}^n} -2\langle \mathbf{z}_{\mathcal{I}}, \mathbf{s} \rangle_{\mathbb{R}} - w \text{ s.t. } \begin{bmatrix} \mathbb{T}[1; \mathbf{y}] & \mathbf{z} \\ \mathbf{z}^H & w \end{bmatrix} \succeq 0, \mathbf{z}_{\mathcal{R}} = 0,$$

where $\mathbb{T}[1; \mathbf{y}]$ is a $n \times n$ Hermitian Toeplitz matrix whose first column is $[1; \mathbf{y}]$, where $[1; \mathbf{y}]$ is the column vector obtained by appending 1 to \mathbf{y} , and where $\mathcal{R} = [n] \setminus \mathcal{I}$ is the position of redundant symbols. This can be equivalently written as

$$\sup_{\mathbf{z} \in \mathbb{C}^n, w} -2\sqrt{w}|\langle \mathbf{z}_{\mathcal{I}}, \mathbf{s} \rangle_{\mathbb{R}}| - w \text{ s.t. } \begin{bmatrix} \mathbb{T}[1; \mathbf{y}] & \mathbf{z} \\ \mathbf{z}^H & 1 \end{bmatrix} \succeq 0, \mathbf{z}_{\mathcal{R}} = 0.$$

We can simply check that the SDP constraints are symmetric, i.e., if \mathbf{z} satisfies the constraint so does $-\mathbf{z}$. Hence, we have

$$\sup_{\mathbf{z} \in \mathbb{C}^n, w} 2\sqrt{w}|\langle \mathbf{z}_{\mathcal{I}}, \mathbf{s} \rangle_{\mathbb{R}}| - w \text{ s.t. } \begin{bmatrix} \mathbb{T}[1; \mathbf{y}] & \mathbf{z} \\ \mathbf{z}^H & 1 \end{bmatrix} \succeq 0, \mathbf{z}_{\mathcal{R}} = 0.$$

Optimizing first with respect to the variable w , which gives the optimal value $|\langle \mathbf{z}_{\mathcal{I}}, \mathbf{s} \rangle_{\mathbb{R}}|^2$, then with respect to the variable \mathbf{z} , and denoting \mathbf{z} with the variable $\mathbf{u} \in \mathbb{C}^m$ in location \mathcal{I} ,

and 0 elsewhere (i.e. $\mathbf{z}_{\mathcal{R}} = 0$), we obtain

$$\sup_{\mathbf{u} \in \mathbb{C}^m} |\langle \mathbf{u}, \mathbf{s} \rangle_{\mathbb{R}}|^2 \text{ s.t. } \begin{bmatrix} \mathbb{T}[1; \mathbf{y}] & \mathbf{z}(\mathbf{u}) \\ \mathbf{z}(\mathbf{u})^H & 1 \end{bmatrix} \succeq 0.$$

Identifying the constrained set by $\mathcal{C}_{\mathcal{I}}$, the dual-optimal value is indeed $\omega(\mathbf{s}, \mathcal{C}_{\mathcal{I}})$. From strong duality, this value corresponds to the primal-optimal value given by $\text{tr}[\mathbf{H}^*]$. Thus, from definition of $\ell(\mathbf{s}, \mathcal{I})$, it is immediately seen that $\ell(\mathbf{s}, \mathcal{I}) = \omega(\mathbf{s}, \mathcal{C}_{\mathcal{I}})/m$. This completes the proof. \blacksquare

Although further analysis is needed to study the behaviour of PAPR loss $\omega(\mathbf{s}, \mathcal{C}_{\mathcal{I}})/m$, our numerical simulations in Section 7 suggest that it concentrates very well around its average $\bar{\omega}(\mathcal{C}_{\mathcal{I}})/m$, where the average seems to be a constant. In [33], it was proved that if a fraction of subcarriers is reserved to reduce the PAPR so that the constant PAPR criterion is always met by *all* the transmitted waveforms, then the fraction of the remaining subcarriers that can be allocated to the information symbols tends to zero. Our numerical results, however, indicate that this seems to be a worst case measure, and in fact the capacity seems to be nonzero. However, to show this rigorously, we need to prove that for any $\lambda \in (0, 1)$, and for $m = n\lambda$,

$$\lim_{n \rightarrow \infty} \mathbb{P}[\bar{\omega}(\mathcal{C}_{\mathcal{I}})/m \leq \mathcal{K}(\lambda)] = 1, \quad (29)$$

for a fixed function $\mathcal{K} : (0, 1) \rightarrow \mathbb{R}_+$. We leave this as a future work to be investigated further.

7 SIMULATION RESULTS

Fig. 3 shows the complementary density function (CCDF) of the random variable $\ell(\mathbf{s}, \mathcal{I})$ for the TR algorithm. We assume that the transmitter uses a 16-QAM (quadrature-amplitude modulation). We compare the results with a case in which the symbols are sampled from a circularly symmetric complex Gaussian distribution. The results show a sharp transition in CCDF for a sufficiently large redundancy (fraction of reserved tones).

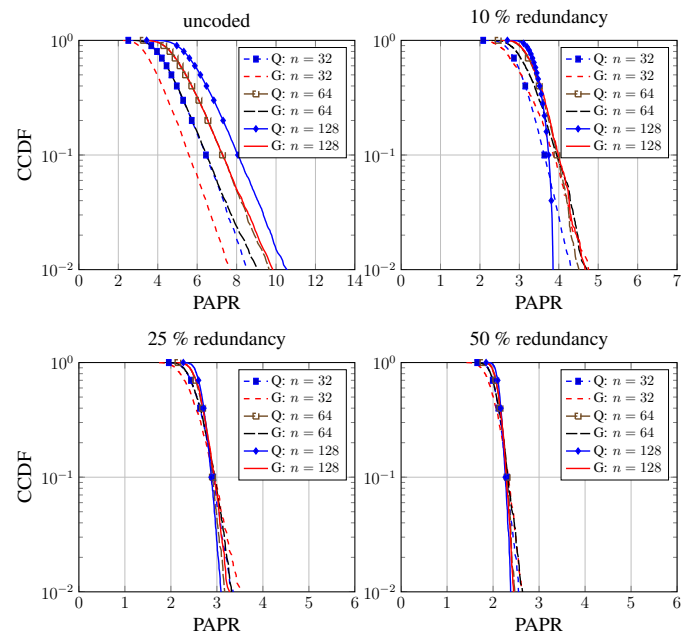


Fig. 3: PAPR performance of QAM (Q) and Gaussian (G) signals as a function of block-length n and redundancy.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Tech. J.*, pp. 379–423, 623–656, July and October 1948.
- [2] D. Slepian, "Some comments on fourier analysis, uncertainty and modeling," *SIAM review*, vol. 25, no. 3, pp. 379–393, 1983.
- [3] J. G. Smith, "The information capacity of amplitude-and variance-constrained scalar gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203–219, 1971.
- [4] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature gaussian channels," *Information Theory, IEEE Transactions on*, vol. 41, no. 4, pp. 1060–1071, 1995.
- [5] A. L. McKellips, "Simple tight bounds on capacity for the peak-limited discrete-time channel," in *IEEE International Symposium on Information Theory*, 2004.
- [6] A. Thangaraj, G. Kramer, and G. Bocerer, "Capacity bounds for discrete-time, amplitude-constrained, additive white gaussian noise channels," *arXiv preprint arXiv:1511.08742*, 2015.
- [7] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.
- [8] B. Farrell and P. Jung, "A kashin approach to the capacity of the discrete amplitude constrained gaussian channel," in *SAMPTA'09*, 2009, pp. Special-Session.
- [9] Y. Lyubarskii and R. Vershynin, "Uncertainty principles and vector quantization," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3491–3501, 2010.
- [10] B. S. Kashin, "Diameters of some finite-dimensional sets and classes of smooth functions," *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 41, no. 2, pp. 334–351, 1977.
- [11] R. v. Nee and R. Prasad, *OFDM for wireless multimedia communications*. Artech House, Inc., 2000.
- [12] G. Wunder, R. F. Fischer, H. Boche, S. Litsyn, and J.-S. No, "The papr problem in ofdm transmission: New directions for a long-lasting problem," *Signal Processing Magazine, IEEE*, vol. 30, no. 6, pp. 130–144, 2013.
- [13] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *Wireless Communications, IEEE*, vol. 12, no. 2, pp. 56–65, 2005.
- [14] J. A. Davis and J. Jedwab, "Peak-to-mean power control and error correction for ofdm transmission using golay sequences and reed-muller codes," *Electronics Letters*, vol. 33, no. 4, pp. 267–268, 1997.
- [15] H. Ochiai, "Block coding scheme based on complementary sequences for multicarrier signals," *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 80, no. 11, pp. 2136–2143, 1997.
- [16] R. D. Van Nee, "Ofdm codes for peak-to-average power reduction and error correction," in *Global Telecommunications Conference, 1996. GLOBECOM'96. Communications: The Key to Global Prosperity*, vol. 1. IEEE, 1996, pp. 740–744.
- [17] K. G. Paterson, "Generalized reed-muller codes and power control in ofdm modulation," *Information Theory, IEEE Transactions on*, vol. 46, no. 1, pp. 104–120, 2000.
- [18] K. G. Paterson and V. Tarokh, "On the existence and construction of good codes with low peak-to-average power ratios," *Information Theory, IEEE Transactions on*, vol. 46, no. 6, pp. 1974–1987, 2000.
- [19] R. You and J. M. Kahn, "Upper-bounding the capacity of optical im/dd channels with multiple-subcarrier modulation and fixed bias using trigonometric moment space method," *Information Theory, IEEE Transactions on*, vol. 48, no. 2, pp. 514–523, 2002.
- [20] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power-and band-limited optical intensity channels corrupted by gaussian noise," *Information Theory, IEEE Transactions on*, vol. 50, no. 5, pp. 784–795, 2004.
- [21] J. Ilic and T. Strohmer, "Papr reduction in ofdm using kashin's representation," in *Signal Processing Advances in Wireless Communications, 2009. SPAWC'09. IEEE 10th Workshop on*. IEEE, 2009, pp. 444–448.
- [22] B. Dumitrescu, *Positive trigonometric polynomials and signal processing applications*. Springer Science & Business Media, 2007.
- [23] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, "Linear matrix inequality formulation of spectral mask constraints with applications to fir filter design," *Signal Processing, IEEE Transactions on*, vol. 50, no. 11, pp. 2702–2715, 2002.
- [24] K. Mahler, "Ein übertragungsprinzip für konvexe körper," *Časopis pro pěstování matematiky a fysiky*, vol. 68, no. 3, pp. 93–102, 1939.
- [25] L. Santaló, "Un invariante afín para los cuerpos convexos del espacio de n dimensiones," *Portugaliae Mathematica*, vol. 8, fasc. 4, p. 155–161, 2009.
- [26] G. Kuperberg, "From the mahler conjecture to gauss linking integrals," *Geometric And Functional Analysis*, vol. 18, no. 3, pp. 870–892, 2008.
- [27] I. Schoenberg, "An isoperimetric inequality for closed curves convex in even-dimensional euclidean spaces," *Acta mathematica*, vol. 91, no. 1, pp. 143–164, 1954.
- [28] C. A. Rogers and G. C. Shephard, "The difference body of a convex body," *Archiv der Mathematik*, vol. 8, no. 3, pp. 220–233, 1957.
- [29] G. Wunder and H. Boche, "Peak value estimation of band-limited signals from their samples with applications to the crest-factor problem in ofdm," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*. IEEE, 2002, p. 17.
- [30] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
- [31] R. Vershynin, "Estimation in high dimensions: a geometric perspective," *arXiv preprint arXiv:1405.5103*, 2014.
- [32] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [33] H. Boche and B. Farrell, "On the peak-to-average power ratio reduction problem for orthogonal transmission schemes," *Internet Mathematics*, vol. 9, no. 2-3, pp. 265–296, 2013.